Novel Tricistronic Vectors and Uses Therefor

This application claims priority to U.S. provisional application serial number 60/399,150 filed July 30, 2002. The entirety of this application is hereby expressly incorporated by reference.

Background of the Invention

Field of the Invention

The field of the invention relates generally to the expression of recombinant DNA. More particularly, the invention relates to novel vectors (and uses therefor) that can be used to express at least three exogenous genes under the control of a single promoter.

Background

A persistent problem associated with expression of multiple, individual recombinant polypeptides (i.e. polypeptides that are not fused to each other) via a vector in an expression system is obtaining satisfactory yields of each polypeptide. This is especially true, for example, when the goal is to express multiple proteins that associate with each other upon expression, where poor yield of one or more of the components will hamper or prevent association of the expressed proteins.

The cloning, transformation and expression efficiencies of a vector typically are inversely related to its size, and therefore one common strategy for expressing multiple polypeptides in an expression system is to use multiple vectors instead of "overloading" a single vector. This approach has drawbacks, however. For instance, short of employing a selection protocol for each vector, there is no way to determine with certainty that a cell contains each vector. In addition, vector incompatibility can hinder obtaining suitable expression levels even where there is satisfactory vector uptake by the cells.

A separate approach is to integrate each exogenous gene into a single construct, but under the control of multiple promoters within that construct. This strategy, too, is riddled with disadvantages. For example, obtaining suitable expression requires successful function of multiple promoters, which can be difficult to achieve. Accordingly, there is no way to determine with certainty that a cell contains sufficient levels of each recombinant polypeptide, short of employing a selection protocol for each gene expression product operatively linked to its respective promoter. Furthermore, utilizing one promoter per exogenous gene disadvantageously results in a relatively large vector. Placing all cistrons into a single vector under the control of a single promoter has not been a viable option in nearly all applications, since, *e.g.*, the further a cistron is positioned from its promoter, the less likely is the chance that acceptable expression yields will be obtained for that cistron.

Certain tricistronic vectors are known in the art, however. For example, Burger *et al.*, *Appl. Microbiol. Biotechnol.* (1999) 52: 345-353 reported a tricistronic vector that encoded , in a 5-prime to 3-prime orientation, (i) a murine light chain Ig, (ii) a murine heavy chain Ig-TNFα fusion and (iii) puromycin acetyltransferase (*pac*) as a selective marker. Burger *et al.* stated that the foregoing tricistronic vector was selected because "expression of the selective marker and product are strictly linked" (*id.* at 351, rt. col.).

However, in Burger *et al.*, the non-Ig polypeptide (*i.e.*, *pac*) functioned only as a selection vehicle and, hence, did not otherwise associate or otherwise interact with either the murine light chain Ig or heavy chain Ig-TNFα fusion. Accordingly, Burger *et al.* provides no suggestion that three "structural" polypeptide domains could be expressed in sufficient yields so as to associate or otherwise interact with each other after expression. In other words, the disclosure by Burger *et al.* did not overcome the prejudice in the art against using a tricistronic vector to express three or more polypeptide domains that associate or otherwise

interact with each other subsequent to expression. It is apparent, therefore, that a vector that satisfies these and other drawbacks known in the art is greatly to be desired. The present invention provides such vectors, together with methods for their use.

Summary of the Invention

Accordingly, it is an object of the invention to provide enhanced expression vehicles for generating at least three polypeptide molecules that can interact with each other subsequent to expression.

It is a further object of the invention to provide enhanced expression vehicles that are compatible with a variety of prokaryotic hosts.

It is still a further object of the invention to provide methods of using the foregoing expression vehicles to discover new and improved therapeutics for treating disease.

These and other objects are made possible with reference to the teachings contained herein.

In one aspect, the invention provides a tricistronic vector construct that comprises a prokaryotic promoter, a first nucleic acid sequence encoding an immunoglobulin-presenting polypeptide, a second nucleic acid sequence encoding a first immunoglobulin (Ig) polypeptide, a third nucleic acid sequence encoding a second Ig polypeptide; a first associating agent fused to or comprised within said Ig-presenting polypeptide, and a second associating agent fused to or comprised within said first Ig polypeptide. The first, second and third nucleic acid sequences are under the control of said promoter and, upon expression of the tricistronic vector, the Ig-presenting polypeptide and the first Ig polypeptide associate via their respective associating agents and the first and second Ig polypeptides self-associate. The vector may optionally be a phagemid vector.

In one embodiment, the Ig-presenting polypeptide may be a phage coat protein, for example, a gIII protein or a functional fragment of a gIII protein. The gIII functional fragment may contain an N-terminal domain of gIII.

In another embodiment, the first and second Ig polypeptides self-associate to form a Fab or other functional Ig fragment, for example via a disulfide bond. The first and/or second associating agent may be a cysteine residue.

In still another embodiment, the first and second Ig polypeptides self-associate via non-covalent interactions.

In other embodiments, the vector contains (i) a first secretory signal sequence in the same reading frame as the nucleic acid sequence encoding the first Ig polypeptide, and/or a second secretory signal sequence in the same reading frame as the nucleic acid sequence encoding the second Ig polypeptide, and/or a third secretory signal sequence in the same reading frame as the nucleic acid sequence encoding the Ig-presenting polypeptide. The first, second and third secretory signal sequences may be prokaryotic signal sequences. The vector may further contain a ribosome binding site positioned 5-primeward of any or all of the nucleic acid sequences encoding the second Ig polypeptide, the first Ig polypeptide and/or the Ig-presenting polypeptide.

In still further embodiments, the associating agents become disassociated in solution upon the addition of a reducing agent. Alternatively, the second associating agent is fused to said first Ig polypeptide via a peptide linker.

The following text provides a more detailed, but non-limiting description of the present invention.

Brief Description of the Figures

Figure 1 is a schematic depiction of principal components of an inventive tricistronic vector, *i.e.*, a single promoter, an Ig-presenting polypeptide, and two Ig polypeptides. Abbreviations: Lac p/o lac promoter operator region; SS gpIII signal sequence, gIII phage gene III; RBS Ribosomal binding site; ompA outer membrane protein A signal sequence; phoA alkaline phosphatase signal sequence; L-His6 PGGSGH6 linker.

Figure 2A is a vector map of an illustrative vector according to the present invention.

Figure 2B provides the nucleic acid sequence for the vector described in Figure 2a.

Figure 3 is a gel that represents a quantitative analysis (by anti-gIIIp Western blot) of the mean display rate of Fab on the surfaces of phage .

Figure 4A is a gel that represents the display rate of a monocistronic scFv vector (pMORPH13) encoding scFvs from a VL-λ pool (conventional display).

Figure 4B is a gel that represents the display rate of a monocistronic scFv vector (pMORPH13) encoding scFvs from a VL-κ pool (conventional display).

Figure 4C is a Vector map for pMorph13 scFv Mac1-5

Figure 4D is the nucleic acid sequence for pMorph13 scFv Mac1-5

Figure 5A is a gel that represents the display rate of a dicistronic scFv vector (pMORPH20) encoding scFvs from a VL-λ pool (display via Cys residues).

Figure 5B is a gel that represents the display rate of a dicistronic scFv vector (pMORPH20) encoding scFvs from a VL-κ pool (display via Cys residues).

Figure 5C is a Vector map for pMorph20 Mac1-5

Figure 5D is the nucleic acid sequence for pMorph20 Mac1-5

Figure 6A is a gel that represents the display rate of a dicistronic Fab vector (pMORPH18) encoding a Fab of framework combination VH2 λ-1; (conventional display).

Figure 6B is a gel that represents the display rate of a dicistronic Fab vector (pMORPH18) encoding a Fab of framework combination VH3 κ-1; (conventional display).

Figure 6C is a Vector map of pMORPH®18-Fab Mac1-5

Figure 6D is the nucleic acid sequence for pMORPH®18-Fab Mac1-5

Figure 7A is a gel that represents the display rate of a dicistronic Fab vector, using a two-vector system (pMORPHX10 & pBR_C_gIII) and encoding a Fab of framework combination VH3 κ-1, respectively (display via Cys residues).

Figure 7B is a gel that represents the display rate of a dicistronic Fab vector, using a two-vector system (pMORPHX10 & pBR_C_gIII) and encoding a Fab of framework combination VH2 κ-1, respectively (display via Cys residues).

Figure 7C is the vector map for pMORPHX10 Fab Mac1-5 VL LHC VH FS

Figure 7D is the nucleic acid sequence for pMORPHX10 Fab Mac1-5 VL LHC VH FS

Figure 7E is the vector map for pMORPHX10 Fab Mac1-5 VL VH LHC

Figure 7F is the nucleic acid sequence for pMORPHX10 Fab Mac1-5 VL VH LHC

Figure 7G is the vector map for pBR-C-gIII

Figure 7H is the nucleic acid sequence for pBR-C-gIII

Figure 8A is a gel that represents the display rate of a tricistronic Fab vector (pMORPH23) encoding a Fab pool (framework combinations VH3 κ/λ).

Figure 8B is a gel that represents the display rate of a tricistronic Fab vector (pMORPH23) encoding a Fab pool (framework combinations VH3 κ/λ).

Figure 9 is a bar graph comparing the functionality and the binding efficiency of Fab-presenting phage of (i) dicistronic Cys display vectors (2-vector system), (ii) tricistronic Cys display vectors, and (iii) dicistronic conventional display vectors in phage ELISA.

Detailed Description

The present invention provides novel tricistronic vectors that are useful in multiple contexts. The inventors surprisingly found that tricistronic vectors may constructed to express three polypeptide molecules in a suitable yield under the control of a single promoter, with the additional feature that the expressed polypeptide domains can maintain function and interact with each other. Another surprising was the observation that all three polypeptides could be exported to the periplasm of a prokaryotic host following expression in the host's cytosol/cytoplasm, and that the expressed polypeptides could interact or otherwise associate in the periplasmic space. Vectors according to the present invention are suitable for use in a number of prokaryotic expression systems.

A. Components of a Vector of the Invention

The components of a tricistronic vector of the present invention include: (i) nucleic acid sequences encoding three polypeptide molecules (non-fused to each other) and (ii) a single promoter that controls expression of all three polypeptides. The polypeptide-encoding nucleic acid sequences encode, for example, (i) an immunoglobulin (Ig)-presenting polypeptide domain, (ii) a first Ig domain, and (iii) a second Ig domain. In addition, a vector of the invention preferably contains a ribosome binding site 5'-ward of each of the foregoing polypeptide molecules, which can enhance expression levels. Upon expression, the two Ig domains associate to form a functional immunoglobulin fragment, which further associates with the Ig-presenting domain, thereby permitting, for example, display of the functional immunoglobulin fragment on the surface of a filamentous phage.

A vector of the invention may optionally contain nucleic acid sequences encoding at least two associating agents, one of which can be fused to (or comprised within) the Ig-presenting polypeptide; and the other of which can be fused to (or comprised within) the first

Ig polypeptide or second Ig polypeptide. Preferably, subsequent to expression of the vector, the Ig-presenting polypeptide and an Ig polypeptide interact with each other via their respective associating agents, and the two Ig polypeptides associate, *e.g.*, by self-association, hydrogen bonding, van der Waals forces, or via an associating agent(s). The foregoing interaction and association interaction and association preferably occur in the periplasm of the prokaryotic host; however, the invention also contemplates association and interaction in the host's cytosol.

a. Promoter:

As used herein, a "promoter" for use in a tricistronic vector of the invention is a promoter that is capable of driving the expression of (i.e. that is functionally linked to) a nucleic acid construct that encodes at least three independent polypeptide molecules (*e.g.*, an Ig-presenting domain and two Ig domains), where those polypeptides are not expressed as fusion proteins with each other. Suitable promoters for use in the invention include, but are not limited to, the lac/operon promoter, CMV promoter P_{bad}, P_{tet}, P_{ara}, P_{ADH1}, P_{GAL}, $P_{EF\text{-}1\alpha}$, P_{SV40}, EM-7 promoter, P_{TEF1}, P_{RSV}, P_{UbC}.

Prokaryotic promoters of the invention can be either constitutive or, more preferably, regulatable (*i.e.*, inducible or derepressible). Further examples of suitable prokaryotic promoters include promoters capable of recognizing the T4 (Malik *et al.*, *J. Biol. Chem.* (1984) 263:1174-1181; Rosenberg *et al.*, *Gene* (1987) 59:191-200; Shinedling *et al.*, *J. Molec. Biol.* (1987) 195:471-480; Hu *et al.*, *Gene* (1986) 42:21-30), T3, Sp6, and T7 (Chamberlin *et al.*, *Nature* (1970) 228:227-231; Bailey *et al.*, *Proc. Natl. Acad. Sci.* (U.S.A.) (1983) 8024:2814-2818; Davanlook *et al.*, *Proc. Natl. Acad. Sci.* (U.S.A.) (1984) 81:2035-2039) polymerases; the P_R and P_L promoters of bacteriophage lambda (THE BACTERIOPHAGE LAMBDA, Hershey, A. D., Ed., Cold Spring Harbor Press, Cold Spring Harbor, N.Y. (1973);

LAMBDA II, Hendrix, R. W., Ed., Cold Spring Harbor Press, Cold Spring Harbor, N.Y. (1980)); the trp, recA, heat shock, and lacZ promoters of *E. coli*; the α-amylase (Ulmanen *et al., J. Bacteriol.* (1985) 162:176-182) and the Σ28-specific promoters of *B. subtilis* (Gilman *et al., Gene* (1984) 32:11-20); the promoters of the bacteriophages of Bacillus (Gryczan, T. J., In: THE MOLECULAR BIOLOGY OF THE BACILLI, Academic Press, Inc., NY (1982)); Streptomyces promoters (Ward *et al., Mol. Gen. Genet.* (1986) 203:468-478); the int promoter of bacteriophage lambda; the bla promoter of the β-lactamase gene of pBR322, and the CAT promoter of the chloramphenicol acetyl transferase gene of pBR325, etc. Prokaryotic promoters are reviewed by Glick, B. R. (*J. Indust. Microbiol.* (1987) 1:277-282); Cenatiempo, Y. (*Biochimie* (1986) 68:505-516); Watson, J. D., (In: MOLECULAR BIOLOGY OF THE GENE, 4th Ed., Benjamin/Cummings Publishing Co., Inc., Menlo Park, Calif. (1987)); Gottesman, S. (*Ann. Rev. Genet.* (1984) 18:415-442)). Other prokaryotic promoters that may be used include other E. coli promoters (Harley *et al., Nucl. Acid Res.* (1987) 15:2343-2361), and Streptomyces promoters (Strohl, *Nucl. Acid Res.* (1992) 20:961-974) for use in Streptomyces species expression hosts. All of the foregoing references are incorporated by reference.

b. Immunoglobulin-presenting polypeptide:

An "immunoglobulin-presenting" or "Ig-presenting" polypeptide or polypeptide domain, as used herein, is a (poly)peptide or protein/polypeptide domain that can interact with at least one immunoglobulin polypeptide, such that the immunoglobulin(s) are able to specifically bind, or are involved in the process of specifically binding, an antigen. An Ig-presenting polypeptide preferably interacts with an Ig domain via an associating moiety that customarily is fused to (or contained within) the Ig-presenting domain.

Suitable Ig-presenting domains include a phage coat (capsid) protein, for example a filamentous phage coat protein. A suitable phage coat protein can be, for example, gene III protein (gIIIp), gene VI protein (gVIp), gene VII protein (gVIIp), gene VIII protein (gVIIIp), and gene IX protein (gIXp). A preferred phage coat protein is gIIIp. A phage coat protein may be either a wild type or a modified protein. A "wild type phage coat protein" refers to any protein forming the phage coat of a naturally occurring bacteriophage. The sequences of the foregoing phage coat proteins (including the differences between the closely related members of the filamentous bacteriophages such as f1, fd, and M13) are well known to those of skill in the art (see, e.g., Kay *et al.*, 1996). The skilled artisan will recognize that other Ig-presenting domains are suitable for use in the present invention.

An Ig-presenting polypeptide of the invention also may be a truncated or modified variant of a phage coat protein (*e.g.*, the C-terminal domain of gIIIp). In this regard, a "truncated" or "modified" variant (or a functional fragment thereof) refers to any phage coat protein that has been modified by deleting, inserting and/or substituting at least part of the wild type sequences. Examples of such variants include truncated gene III protein variants that have been found in bacteriophage mutants (see, for example, Crissman & Smith, 1984) or that have been generated for use in phage display methods (e.g. Bass *et al.*, 1990; Krebber, 1996).

The invention also contemplates the use of other Ig-presenting polypeptides. An Ig-presenting polypeptide also may be a green fluorescent protein (gfp), any protein of the cell surface or of the cell wall of bacterial cell, or any protein of a bacteriophage or virus coat.

c. Immunoglobulin or "Ig" polypeptide or domain

An "immunoglobulin" or "Ig" polypeptide or domain hereby is defined as a domain of the protein class IgG, IgM, IgE, IgA, and IgD (and any subclass thereof), and includes all

conventionally known antibodies and functional fragments thereof. A "functional fragment" refers to a fragment of an immunoglobulin which retains the antigen-binding moiety of an immunoglobulin. A preferred class of immunoglobulins for use in the present invention is IgG. More specifically, an immunoglobulin domain of the invention can include the domain of (i) a $F(ab')_2$, fragment, or (ii) a Fab fragment. The $F(ab')_2$, or Fab may be engineered to minimize the intermolecular disulphide interactions that occur between the C_{H1} and C_L domains. An Ig polypeptide may have an amino acid sequence derived from that of an antibody isolated from nature or derived from a natural source, or may have a sequence that is designed *in silico* and encoded by a nucleic acid that is synthetically created. *In silico* design of an antibody sequence can be achieved, for example, by analyzing a database of sequences and devising a polypeptide sequence utilizing the data obtained therefrom. Methods for designing and obtaining such *in silico*-created sequences are described, for example, in U.S. Patent No. 6,300,064 to Knappik *et al.*, which hereby is incorporated by reference in its entirety.

A tricistronic vector of the invention preferably encodes two Ig polypeptides that interact with each other and form a functional (antigen-binding) molecule. Interaction between the two Ig polypeptides typically is mediated by residues that belong to each Ig polypeptide. To this end, the first and second polypeptide can comprise heavy and light chain regions of an antibody that associate via non-covalent interactions between corresponding heavy and light chain domains, such as between VH and VL in an Fv fragment, or between VH/VL and CH1/CL in a Fab fragment. Additionally, heavy and light chain regions of an antibody may associate by forming disulphide bonds between the two chains, such as is possible in a Fab fragment. The present invention specifically contemplates the interaction of two Ig polypeptides by mechanisms other than formation of one or more

inter-chain disulfide bonds , *e.g.*, via a linker moiety that is non-covalently attached to at least one of the Ig domains, via hydrogen bonds, via van der Waals interactions, or via peptidic association domains fused to the Ig polypeptides, such as described in U.S. Patent No. 6,294,353 to Pack et al., which hereby is incorporated by reference.

An "associating agent" for use in the present invention is defined as an agent that can bring about the interaction between expressed Ig-presenting and Ig polypeptides. An expressed associating agent of the invention is fused to, or comprised within, an Ig-presenting polypeptide and a complementary associating agent is fused to, or comprised within, an Ig polypeptide. The foregoing associating agents may be two different agents, or may be two identical or substantially identical agents. An associating agent according to the invention preferably contains a cysteine residue that is available for the formation of an intermolecular disulphide linkage.

Preferably, the associating agents are selected so that they do not interfere with the desired function of the fully associated protein complex. Typically, therefore, the associating agents are suitable amino acid residues that are located outside the region(s) deemed to be responsible for a putative function of the (poly)peptide/protein of interest such as binding to a target. For example, a cysteine residue that is intended to form an inter-chain disulfide bond is positioned at, or in the vicinity of, either the N- or the C-terminus of a polypeptide.

Other suitable associating agents include those which (i) can be fused to the C-terminal end of an Ig polypeptide (or within about 15 amino acid residues thereof) and (ii) can interact with an associating agent fused to or comprised within an Ig-presenting polypeptide. Likewise, suitable agents include any which (i) can be fused N-terminally to (or comprised within) an Ig-presenting polypeptide (*e.g.*, phage coat protein) and (ii) can interact

with an associating agent fused to or comprised within an Ig polypeptide. A specific example of a pair of associating agents in this regard is an avidin-biotin complex..

In the context of the present invention, a cysteine residue is "available for the formation of an intermolecular disulfide bond" if the residue is (i) located N-terminal, C-terminal, or internal to a polypeptide and (ii) accessible for the formation of a disulfide bond with a second residue of the same or different kind. This includes cysteine residues that are buried, and thus not accessible in the "final" polypeptide molecule, but which are accessible in an intermediate compound formed in the course of expression, processing and/or transport in a host cell.

In one embodiment, two associating agents may associate, or attach, by the formation of a disulfide bond between (i) at least one cysteine residue present in an Ig polypeptide and (ii) a second cysteine residue present within an Ig-presenting domain that is a wild type phage coat protein. In the case of filamentous bacteriophage fd, for example, wild type proteins contain the following cysteine residues: Cys7, Cys36, Cys46, Cys53, Cys188, Cys201, Cys354, and Cys371 of protein III; residue Cys84 of protein VI; residue Cys22 of protein VII; residue Cys16 of protein IX. Any one or more of these residues may act as an associating agent.

A tricistronic vector of the invention also may contain one or more ribosome binding sites. A ribosomal binding site (Shine-Dalgarno sequence) is a purine rich sequence that in on bacterial mRNA is located about ten nucleotides 5-primeward of the initiator codon for a particular polypeptide. A Shine-Dalgarno sequence is involved in the binding of the ribosome and the mediating of efficient translation of the respective gene.

A tricistronic vector of the invention also may contain one or more nucleic acid sequences that encode a signal or secretory polypeptide. A "signal" or "secretory"

polypeptide hereby is defined as a polypeptide responsible for transporting another polypeptide from bacterial cytosol to bacterial periplasm. A signal or secretory polypeptide of the invention preferably is located N-terminal to the polypeptide to be transported to the periplasm. The use of one or more secretory polypeptides can be especially advantageous in the context of phage display technology, as described, *infra*, whereby the secretory polypeptide (i) is linked to a encoded polypeptide, and (ii) directs the corresponding polypeptide to the periplasmic space of its prokaryotic host cell. Secretory polypeptides include, for example, ompA and phoA, gene III signal sequence, st II, and pelB, each of which can be used in a prokaryotic expression system. Other nucleic acid sequences encoding secretory peptide sequences are well known in the art and may also be used in the present invention. In one aspect of the invention, a secretory nucleic acid sequence (*e.g.*, ompA) is linked to the nucleic acid sequence that encodes a first Ig domain, while a second secretory nucleic acid sequence (*e.g.*, phoA) is linked to the nucleic acid sequence that encodes a second Ig domain. A secretory nucleic acid sequence also can be linked to the nucleic acid sequence that encodes an Ig-presenting polypeptide. Alternatively, the secretory domain can be an inherent property of an Ig-presenting domain of the invention.

A tricistronic vector of the invention also may contain one or more nucleic acid sequences that can encode a "polypeptide linker" that functions to link an associating agent to an Ig-presenting and/or an Ig domain. In this context, the linker can be viewed as a "spacer" between an associating agent and its respective polypeptide. This linker preferably contains about 1-50 amino acids, and preferably 5 to 15 amino acids. Typically, a linker consists of glycine-serine rich stretches, but can also contain other amino acid residues. The size can also be variable according to the purpose of the linker.

A tricistronic vector of the invention also can be constructed so as to contain one or more affinity tags (e.g., His6 tag) that is fused to one of the Ig domains, for example. An affinity tag can be used to purify or isolate a population of Ig molecules bearing this tag.

A tricistronic vector of the invention also can be constructed so as to contain one or more restriction sites that facilitate cloning, sub-cloning, or other manipulation of the vector. For example, when a plurality of restriction sites are present, unique restriction sites can be engineered to flank a particular segment of the vector, thereby making the vector modular. The feature of modularity can be advantageous, *e.g.*, for subsequent modification of the tricistronic vector at one or more discrete positions. According to this approach, a particular segment of the tricistronic vector can be excised and substituted with another desired segment, using convention technology. A library such as the HuCAL antibody library described in U.S. Patent No. 6,300,064 to Knappik *et al.*, is particularly preferred for use in a vector of the present invention.

An illustrative, non-limiting embodiment of a vector according to the invention (pMORPH23) is set forth in Figure 2A. According to Figure 2A, pMORPH23 contains a ColEI origin of replication, a functional origin for single stranded replication, and a chloramphenicol-resistance gene. The tricistronic operon is under the control of an inducible *lac* promoter/operator region. All functional modules are flanked by unique restriction sites. The first expression cassette contains the signal sequence of geneIII, and the engineered full-length (mature) geneIII sequence with an additional N-terminal cysteine residue. The second expression cassette, which is preceded by a ribosomal binding site (SD-Seq), encodes the light chain of an Ig and contains the bacterial signal sequence ompA followed by VL and CL. The third expression cassette, which is preceded by a ribosomal binding site (SD-Seq), contains a heavy chain Fd (VH1 + CH1) with an additional C-terminal cysteine. The

bacterial signal sequence phoA is followed by VH1 and CH1, whereby a glycine/serine-rich linker and a His6-tag act as a spacer for the introduced cysteine to the Fd chain..

B. Constructing a Vector of the Invention

Methods for constructing vectors comprising nucleic acid molecules are known in the art (*see, e.g.*, Sambrook *et al.*, 1989; Ausubel *et al.*, 1994). A vector map of a representative vector of the invention (pMORPH23) is provided in Figure 2A, with its nucleic acid sequence provided in Figure 2B.

C. Representative Uses of a Vector According to the Invention

A tricistronic vector of the invention can be used, or can be modified to be used, in a variety of prokaryotic expression systems. A suitable host cell is any cell that permits expression and subsequent interaction of the three principal polypeptide domains (*i.e.*, Ig-presenting and two Ig domains). Methods for introducing vectors into appropriately chosen host cells, and causing or allowing the expression of polypeptides are known in the art (*see, e.g.*, Sambrook *et al.*, 1989; Ausubel *et al.*, 1994).

A vector according to the invention, is particularly suited for expression in an *E.coli* host cell. In this regard, the vector can be in the form of a phagemid vector. A phagemid consists of elements of conventional plasmid vectors (*e.g.*, marker gene, cloned genes, plasmid origin of replication) and of elements of filamentous phage (*e.g.*, gIII, PS and phage *ori*). A phagemid can be introduced into a host cell, and subsequently be cultivated and amplified therein like a plasmid. Phagemid vectors are well known in the art.

A phagemid does not encode all of the genes necessary to permit assembly of viral particles and requires "rescue" in a host cell with a helper phage. The helper phage provides the missing phage genes that permit assembly of the viral particles. It will be appreciated that use of a phagemid/helper phage system an result in production of particles that contain either

the helper phage genome or the phagemid. Methods of preferentially packaging the phagemid are well known in the art, for example by using a helper phage, such as M13 K07 that contains a functional, but defective, DNA origin of replication so that phagemid is preferentially packaged into phagemid particles. Methods for the introduction of genetic material required to produce progeny phage or phagemid particles in appropriate host cells, and for causing or allowing the generation of such particles are well known in the art (*see, e.g.*, Kay *et al.*, eds. (1996) PHAGE DISPLAY OF PEPTIDES AND PROTEINS: A LABORATORY MANUAL. Academic Press, Inc., San Diego).

A vector of the invention can, accordingly, be use to carry out a method for producing a polypeptide or protein having a desired property. This method includes the steps of (i) providing a collection of bacteriophage particles that present on their surface a diverse collection of one or more Ig polypeptides as defined herein; and (ii) screening and/or selecting the diverse collection for at least one Ig domain having the desired property. Here, the term "desired property" refers to a property that (a) one of the polypeptides or proteins out of the diverse collection should have and (b) forms the basis for screening and/or selecting the diverse collection. A property might be the ability to: bind a target, block a target, or activate a target-mediated reaction. A further property may be, for example, enzymatic activity, or any other properties known to those skilled in the art. Methods for identifying suitable experimental formats and for carrying out necessary steps for performing screening and/or selection are well known in the art.

A preferred property of an Ig polypeptide is specific binding to a target. The target can be presented to the diverse collection of bacteriophage particles in a variety of ways well known to one of ordinary skill, for example, by coating on surfaces for solid phase

biopanning, by linkage to particles such as magnetic beads for biopanning in solution, or by display on the surface of cells for whole cell biopanning.

Bacteriophage particles that display (via an Ig-presenting domain) one or more Ig polypeptides (which are bound to a target) can be recovered by a variety of methods well known to one of ordinary skill. If the associating agents link the Ig-presenting polypeptide and Ig polypeptide via a disulfide bond, then the specifically bound Ig-target complexes can be treated under reducing conditions (*e.g.*, incubation with DTT) to cleave the disulfide bonds and to recover the specific bacteriophage particles for further rounds of biopanning and/or for identification of the Ig polypeptide domains specifically binding to said target.

Examples

The present invention can be better understood with reference to the following examples, which are not intended to limit the scope of the invention as described above.

Example 1: General protocol for quantitative analysis of display of antibody fragments on phage

The protocol, which also applies to Examples 2-6, was performed according to Johansen, L. K. *et al.* (1995), *Protein Engng.* 8, 10, 1063-1067. Different dilutions of the same phage preparation were subjected to a protein gel. However, in Examples 3, 5 and 7, no reducing agents were added, due to the presence of cysteines as associating agents. The proteins of the protein gel were transferred to a membrane. gIIIp protein on the membrane was detected by anti-gIIIp antibody (Western blot). Then, one wild-type gene III protein ("wtgIIIp") band and one band of the antibody-gIIIp linkage, which have the same intensity were analysed. Given (i) the number of phages loaded, (ii) the molar ratio of both proteins, and (iii) the assumption of 5 wtgIIIp-proteins per phage, the mean number of antibody fragments displayed per phage could be calculated. Figure 1 provides expression data of a dicistronic Fab vector (pMORPH18) using conventional (*i.e.*, gIIIp-fusion) display. The data

indicate a ratio of Fd-gIIIct:wtgIIIp = $1.25x10^8$ phage:$2x10^7$ phage. This correlates to the presentation of 1 Fd-gIIIct per 6.25 wtgIIIp; in other words, 1 Fd-gIIIct per 1.25 wtgIIIp. Accordingly, the mean number of Fabs per phage in this experiment was 0.8.

Abbreviations throughout: Fd = VH-CH1; Fd-ct = Fd-gIIIct & VH-CH1-gIIIct; g3p = gIIIp.

Example 2: Display determination of monocistronic scFv vector, using conventional display

A protocol as disclosed in Example 1 was carried out for performing a quantitative display analysis of a monocistronic scFv vector (pMORPH13), using conventional phage display. Figure 4A provides expression data of the (pMORPH13) vector from a VL-λ pool; and Figure 4B provides expression data of the (pMORPH13) vector from a VL-κ pool. The data indicate a ratio of scFv-gIIIct:wtgIIIp = $1x10^9$ phage:$6.7x10^7$ phage in Figure 4A, and a ratio of scFv-gIIIct:wtgIIIp = $5x10^8$ phage:$1x10^7$ phage in Figure 4B. Accordingly, the mean number of ScFv per phage in this experiment was approximately 0.3 and 0.1, respectively.

Example 3: Display determination of a dicistronic scFv vector, using Cys display

A protocol as disclosed in Example 1 was carried out (except with using reducing agents) for performing a quantitative display analysis of a dicistronic scFv vector (pMORPH20), using Cys display. Figure 5A provides expression data of the pMORPH20 vector from a VL-λ pool; and Figure 5B provides expression data of the pMORPH20 vector from a VL-κ pool. The data indicate a ratio of scFv-SS-gIIIct:wtgIIIp = $1x10^{10}$ phage:$4x10^7$ phage in Figure 5A, and a ratio of scFv-SS-gIIIct:wtgIIIp = $5x10^{9***}$ phage:$2x10^7$ phage in Figure 5B. Accordingly, the mean number of scFv per phage in this experiment was approximately 0.02 and 0.02 respectively.

Example 4: Display determination of a dicistronic Fab vector, using conventional display

A protocol as disclosed in Example 1 was carried out for performing a quantitative display analysis of a dicistronic Fab vector (pMORPH18), using conventional phage display. Figure 6A provides expression data of the pMORPH18 vector (single Fab of framework combination VH2-λ1); and Figure 6B provides expression data of the pMORPH18 vector (single Fab of framework combination VH3-κ1). The data indicate a ratio of Fd-gIIIct:wtgIIIp = $1.x10^9$ phage:$2x10^7$ phage in Figure 6A, and a ratio of Fd-gIIIct:wtgIIIp = $1x10^8$ phage:$1x10^7$ phage in Figure 6B. Accordingly, the mean number of Fabs per phage in this experiment was approximately 0.1 and 0.5, respectively.

Example 5: Display determination of a dicistronic Fab two-vector system, using Cys display

A protocol as disclosed in Example 1 was carried out (except with using reducing agents) for performing a quantitative display analysis of a dicistronic Fab vector in a two-vector system (pMORPH10 + pBR_C_gIII), using Cys display. Figure 7A provides expression data of the pMORPH10 vector system (single Fab of framework combination VH3-κ1); and Figure 7B provides expression data of the pMORPH10 vector system (single Fab of framework combination VH2-λ1). The data indicate a ratio of VL_CL-SS-gIII:wtgIIIp = $1x10^9$ phage:$8x10^6$ phage in Figure 7A, and a ratio of VL_CL-SS-gIII:wtgIIIp = $8x10^9$ phage:$3x10^7$ phage in Figure 7B. Accordingly, the mean number of Fabs per phage in this experiment was approximately 0.04 and 0.02, respectively.

Example 6: Analysis of display rates and efficiency in phage ELISAs

The following table summarizes approximate display rates obtained in Examples 1-5:

Table I

System	Vector(s)	Approximate display rates (Ig per Phage)
scFv conventional (monocistronic)	pMORPH13	0.1 - 0.3
scFv CysDisplay (dicistronic)	pMORPH20	0.02
Fab conventional (dicistronic)	pMORPH18	0.1 - 0.8
Fab CysDisplay (dicistronic)	pMORPHX10 + pBR_C_gIII	0.02 - 0.04

From this table, two trends are understood. First, the display rates decrease as much as 2.5 to 40 fold when using CysDisplay in lieu of conventional display. Second, the display rates decrease as much as 5 to 15 fold when moving from a monocistronic conventional display vector to a dicistronic CysDisplay vector. Accordingly, CysDisplay phage generally showed reduced display rates in comparison to phage containing conventional genetic fusions of antibody fragments to gIII (or gIIIct fragment). Because it would be highly undesirable to work with display rates lower than 0.04 Fabs per phage, the use of a tricistronic Fab vector that additionally was engineered for CysDisplay was thought not to be possible, based on the foregoing trends of decreased display rates.

Example 7: Display determination of tricistronic Fab vector, using Cys display (single vector system)

A protocol as disclosed in Example 1 was carried out (except with using reducing agents) for performing a quantitative display analysis of a tricistronic Fab vector (pMORPH23), using Cys display. Figure 8A provides expression data of the pMORPH23 vector from a VH3 + κ/λ pool; and Figure 8B provides expression data of the pMORPH23 vector from a VH3 + κ/λ pool. The data indicate a ratio of Fd-SS-gIII:wtgIIIp = $5x10^{10}$ phage:$3x10^{9}$ phage in Figure 8A, and a ratio of Fd-SS-gIII:wtgIIIp = $5x10^{9}$ phage:$1x10^{8}$ phage in Figure 8B. Accordingly, the mean number of Fabs per phage in this experiment was approximately 0.3 and 0.1, respectively.

As is shown in Table II below, the tricistronic Fab CysDisplay vector (pMORPH23) yield improved Fab display rates when compared to the dicistronic Fab CysDisplay system, which always needs a second vector providing the Cys-gIII construct. When using a constant amount of phage in the foregoing examples, the signals obtained with the tricistronic system were higher than those obtained with the dicistronic system. This indicates an increased display rate with the tricistronic version, which was unexpected.

Table II

System	Vector(s)	Approximate display rates (Ig per Phage)
scFv conventional	pMORPH13	0.1 - 0.3
scFv CysDisplay	pMORPH20	0.02
Fab conventional	pMORPH18	0.1 - 0.8
Fab CysDisplay dicistronic	pMORPHX10 + pBR_C_gIII	0.02 - 0.04
Fab CysDisplay tricistronic	pMORPH23	0.05-0.3

Example 8: Comparison of dicistronic and tricistronic Fab Cys Display vectors in phage ELISA

Phage preparations (i) anti-Mac1 I-domain, (ii) Fab Mac1-5 and (iii) Mac1_A8 were expressed from the dicistronic CysDisplay vector (pMORPHX10 + pBR_C_gIII; two-vector-system), the tricistronic CysDisplay vector (pMORPH23) and the dicistronic, conventional Fab vector (pMORPH18) and displayed on phage. The phage were concentrated and the titer of the phage preparations was determined.

Maxisorp wells of a microtiter plate were coated with 100 µl Mac1 I-domain protein per well (concentration of the antigen solution in PBS: 50 µg/ml) overnight at 4°C. The antigen solution was removed and the coated wells were washed with PBS. Next, the antigen-coated wells were blocked with 300 µl 5% MPBST for 1 hour at room temperature. At the same time, an aliquot of each phage preparation (100 µl per well; 7.5E+9 phages) 1:1 was mixed with 10% MPBST (incl. 0.1% Tween20). The phage were incubated for 1 hour at

room temperature. The coated wells were washed 3x with PBS. Then, 200 µl of pre-blocked phage solution was transferred into each coated well, and incubated for 1 hour at room temperature. Then, the phage were removed from the wells, and non-bound phage were washed off using PBST and PBS. Next, 100 µl anti-M13-HRP conjugate (1:5000) in 5% MPBST (incl. 0.05% Tween20) was added and incubated for one hour at room temperature. Another PBST and PBS wash was performed, and 100 µl POD-Substrate was added. A measurement at 370 nm was taken after several minutes in order to quantify the amount of anti-Mac1 phage attached to the antigen in the wells.

Figure 9 is a bar graph that compares functionality and binding efficiency (functional Fab display) between dicistronic and tricistronic Fab Cys Display vectors in phage ELISA. The improved binding efficiency of the phage resulting from the tricistronic CysDisplay vector versus the dicistronic CysDisplay vector confirms the data of the increased display rates. Bars 1 and 2 represent independent experiments of the same construct. The first two bars for each group represent experiments performed with Fab molecule Mac1-5; the last two bars for each group represent experiments performed with Fab molecule Mac1_A8).

Example 9: Successful Antibody Library Screening using a tricistronic vector system in Cys Display

Wells of MaxiSorp™ microtiter plates (NUNC) were coated with 15 µg per antigen (ICAM-1 protein, rabbit myosin, FITC-BSA, estradiol-BSA) dissolved in PBS. Using a tricistronic vector as described in Examples 2A and 2B in conjunction with proprietary MorphoSys phage display and selection techniques, the results provided in Table III were obtained upon screening a MorphoSys HuCAL® Library.

Table III

Antigen	Elution	% primary hits 2nd round	% primary hits 3rd round	No. of consolidated, specific antibodies
ICAM-1 protein	DTT	0%	17%	1
ICAM-1 protein	glycine + TG1	0%	60%	3
myosin	DTT	1%	29%	4
myosin	glycine + TG1	14%	19%	1
FITC-BSA	DTT	82%	100%	6
FITC-BSA	glycine + TG1	92%	87%	6
estradiol-BSA	DTT	75%	67%	6
estradiol-BSA	glycine + TG1	59%	67%	3

The foregoing data confirm that tricistronic vectors of the invention are effective vehicles for expressing, at a minimum, three functional polypeptide molecules.

Example 10: Construction of pMORPH23 vector

The vector pMORPH23 described here is a derivative of the pCAL vector series (WO 97/08320; Knappik *et al.*, 2000), which is a modified version of the dicistronic expression vector pMORPH20 (example 3). A vector map for pMORPH20 is provided in Figure 5C and the related nucleic acid sequence is provided in Figure 5D.

The dicistronic expression vector pMORPH20 was digested with restriction enzymes Stu*I* and Msc*I*, to remove the scFv–expression module. The resulting blunt end cut vector was religated after agarose gel purification and transformed into competent E.coli cells. The intermediate vector product was further modified by replacing the ompA signal sequence (Xba*I* and Eco*RV* digest) by a oligonucleotide cassette preformed by annealing primer pairs A and B coding for the gpIII signal sequence and introducing a 5' AccI restriction site and a 3' blunt end.

Primer A:

ctagtatacg agggcaaaaa atgaaaaaac tgctgttcgc

gattccgctg gtggtgccgt tctatagcca tagcgactac tgcgac

Primer B:

gtcgcagtag tcgctatggc tatagaacgg caccaccagc

ggaatcgcga acagcagttt tttcattttt tgccctcgta ta

To obtain the final pMORPH23 library cloning vector, an AflII-XbaI-Bla-EcoRI stuffer cassette was introduced by ligation. The stuffer fragment allows efficient cloning of HuCAL Fab fragments by XbaI and EcoRI. An example for tricistronic pMORPH23-HuCALFab vector is shown in Figure 2A and 2B. All three modules in pMORPH23 are transcribed as one unit from the lac p/o region.